

EXHIBIT A

D3

RAPID COMMUNICATION

Salmon Pancreas Disease Virus, an Alphavirus Infecting Farmed Atlantic Salmon, *Salmo salar* L.

Jonathan H. Weston,* Michael D. Welsh,† Marlan F. McLoughlin,† and Daniel Todd†¹

*Department of Veterinary Science, The Queen's University of Belfast, Stormont, Belfast BT4 3SD, United Kingdom; and †Virology Section, Veterinary Sciences Division, Department of Agriculture for Northern Ireland, Stormont, Belfast BT4 3SD, United Kingdom

Received December 9, 1998, returned to author for revision January 19, 1999, accepted February 11, 1999

A 5.2-kb region at the 3' terminus of the salmon pancreas disease virus (SPDV) RNA genome has been cloned and sequenced. The nucleotide and predicted amino acid sequences show that SPDV shares considerable organizational and sequence identity to members of the genus alphavirus within the family *Togaviridae*. The SPDV structural proteins encoded by the 5.2-kb region contain a number of unique features when compared to other sequenced alphaviruses. Based on cleavage site homologies, the predicted sizes of the SPDV envelope glycoproteins E2 (438 aa) and E1 (461 aa) are larger than those of other alphaviruses, while the predicted size of the alphavirus 6K protein is 3.2 K (32 aa) in SPDV. The E2 and E1 proteins each carry one putative N-linked glycosylation site, with the site in E1 being found at a unique position. From amino acid sequence comparisons of the SPDV structural region with sequenced alphaviruses overall homology is uniform, ranging from 32 to 33%. While nucleotide sequence analysis of the 26S RNA junction region shows that SPDV is similar to other alphaviruses, analysis of the 3'-nontranslated region reveals that SPDV shows divergence in this region. © 1999 Academic Press

Introduction. Pancreas disease (PD) of farmed Atlantic salmon, first described in Scotland in 1984 (18), has since been reported throughout Western Europe and North America (11, 24). The disease causes major economic losses with up to 50% mortality in first year salmon smolts in Ireland (29). Postsmolt salmon affected with PD appear runted and characteristic histological lesions in the pancreas, heart, and muscle are observed. The first cell culture isolate of a virus, designated salmon pancreas disease virus (SPDV) and later shown to be the causal agent of the disease (17), was made in this laboratory from diseased fish (19). On the basis of physiochemical characteristics and structural morphology, SPDV was tentatively classified as a "toga-like" virus (19).

In this paper we report the production of cDNA clones specific to SPDV RNA and, on the basis of the sequence analysis of the 5.2-kb region at the 3' terminus of the SPDV genome, we have identified SPDV as an alphavirus, the first reported in fish. The alphaviruses described to date are arthropod-borne viruses that contain a positive sense, single-stranded genome of approximately 12 kb. The four nonstructural proteins (nsP1–4) that are responsible for replicating the virus RNA are encoded in

the 5'-terminal two thirds of the genome. Virus replication involves the synthesis of a subgenomic (26S) RNA species of approximately 4 kb that corresponds approximately to the 3' third of the genome. This subgenomic RNA is translated into a polyprotein, approximately 130 kDa in size, which is cleaved proteolytically into the capsid protein and envelope glycoproteins E2 and E1 (28).

The sequence analysis, reported in this paper, compares the structural protein gene region of the SPDV genome with those of other sequenced alphaviruses.

Results. Cloning strategy. Sequence analyses indicated that 90% of cDNA clones prepared from RNA extracted from gradient-purified SPDV exhibited strong protein homologies to previously sequenced alphaviruses. Three of these clones, N11, N38, and N50, mapped to the 5.2-kb region at the 3' terminus of the consensus alphavirus genome and knowledge of the sequences contained by these clones was used to produce three overlapping cDNA clones, pBS1150, pBS5038, and pBS38R. Clone pBS38R was produced using a 3' RACE reaction, the success of which depended on the SPDV genome possessing a polyadenylated tract at its 3' terminus (Fig. 1). Clone pBS1150 is approximately 2.4 kb and encodes the C-terminus of nsP4, the capsid, E3, and the N-terminus of E2. Clone pBS5038 is approximately 1.6 kb and encodes the C-terminus of E2, 6K, and the N-terminus of E1 and pBS38R is approximately 1.1 kb and

The nucleotide sequence reported in this paper has the GenBank Accession No. AJ012631.

¹ To whom correspondence and reprint requests should be addressed. Fax: +44 (0)1232 525773. E-mail: daniel.todd@dani.gov.uk.



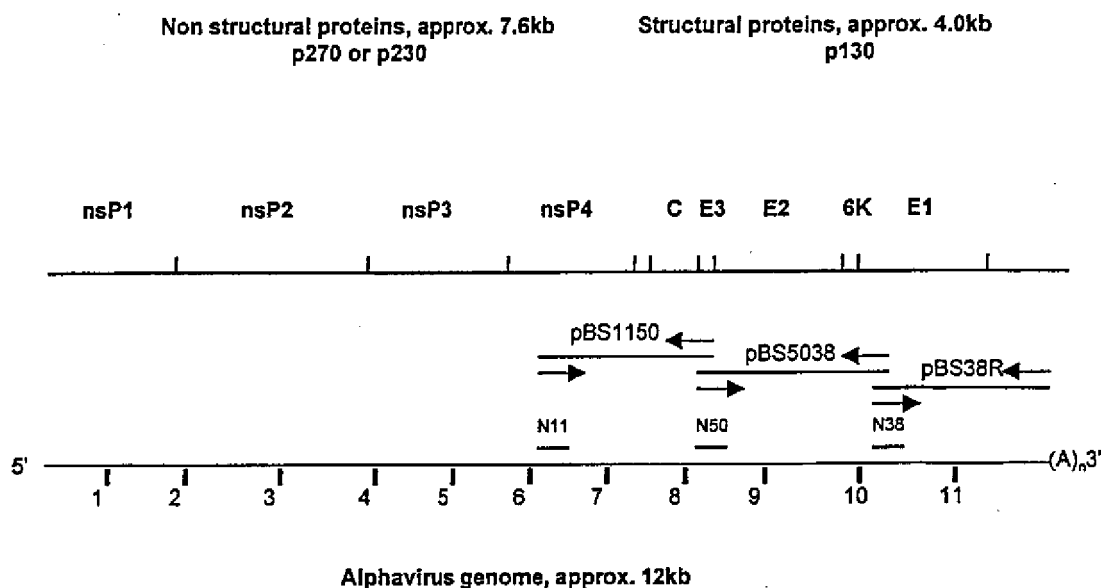


FIG. 1. Cloning and sequencing strategy for SPDV. The position where the clones N11, N50, and N38 map to the 3' terminus of the alphavirus genome is shown. Sequence information from these clones was then used to produce pBS1150, pBS5038, and pBS38R. Sequencing was carried out by primer walking from the ends of each of these clones as indicated by the arrows.

encodes the C-terminus of E1, the 3'-nontranslated region, and the poly(A) tract (Fig. 1).

Nucleotide Sequence of the 3'-Terminal 5.2-kb Region of SPDV Genome. The complete translated nucleotide sequence of the 3'-terminal 5.2-kb region of SPDV RNA is presented in Fig. 2. The genome organization of the virus is similar to that of other alphaviruses (Fig. 1). Following the stop codon of the putative RNA-dependent RNA polymerase, nsP4, there is a nontranslated region of 35 nucleotides followed by an open reading frame of 3852 nucleotides encoding the structural proteins. A 3'-nontranslated region of 91 nucleotides, including the stop codon, was identified prior to the poly(A) tract. The start of the SPDV 26S RNA was identified by alignment to other alphaviruses. The 26S RNA is 3972 nucleotides excluding the poly(A) tract and is shorter than most described alphaviruses.

Comparison of SPDV Structural Proteins with Other Alphavirus Proteins. The putative cleavage sites between the structural proteins of SPDV (Fig. 2) were predicted on the basis of homologies shared with other alphaviruses (data not shown). Cleavage at these sites would result in SPDV having larger E1 and E2 envelope glycoproteins than any previously reported alphavirus and more notably a smaller 6K protein (Table 1). The percentage of amino acid sequence identity was determined using pairwise comparisons between the structural region of SPDV and that of other alphaviruses and was shown to be uniform, with percentage identities falling between 32 and 33% (Table 2). The C-terminal 394 amino acids of nsP4 of

SPDV show 59–62% sequence identity to the equivalent region in other alphaviruses (data not shown).

Comparison of the putative glycosylation sites in the envelope proteins shows SPDV to be unusual in having no N-linked glycosylation sites in E3. In alphaviruses such as SFV and SIN, E3 has been shown to possess a conserved site at amino acids 11–14. The E2 and E1 glycoproteins of SPDV each contain one site. The N-linked site within E2 is at amino acid 319, which is similar to that found in SIN, WEEV, and EEEV. However, within E1 of SPDV, the N-linked site is at amino acid 35, a finding which contrasts with that observed in other sequenced alphaviruses, which are glycosylated at amino acids 139–141.

With most alphaviruses, a major antigenic region is located between amino acids 170 and 220 within the E2 glycoprotein and is responsible for virus neutralization (27, 12). Examination of the sequence of amino acids from 194 to 234 within the E2 specified by SPDV shows some sequence similarity to the antigenic regions of other alphaviruses in that it contains two cysteine residues and one glycine residue which are invariant in this region among alphaviruses (26).

Like other alphaviruses the amino acid sequence of the N-terminus of the capsid protein contains a relatively high proportion of basic residues, which are believed to have a role in the interaction with genomic RNA during encapsidation (8, 5, 22). The SPDV capsid also contains the C-terminus serine protease motif responsible for release of the capsid protein from the structural polyprotein (10, 3). The SPDV E1 glycoprotein contains the con-

ACTATGGACTCAGGGCAATGAACGTGGAGGCTTTTAAAGTTTCGCTGTAGGACACCGACCTGTGGACTGAGTTCGGGAAAAACCA (90)
 T M D S A A M N V E A F K S F A C K D T D L W T E F A E K P
 GTAAGGTTGTGCGCCGGCCAAATCGAAGATATGTCTTTCATCTACAAGGGCCCAAGGCCAATGTGATGCACAGCAGAGTCGAAGCCGTA (180)
 V R L S P G Q I E E Y V F H L Q G A K A N V M H S R V E A V
 TGCCTTGACCTCTCGGAGGTGGCTATGGACAGGTTTCACTAGACATGAACCGCGACGTCRAAGTGACGCCAGGCACGAAGCACCTAGAG (270)
 C P D L S E V A M D R F T L D M K R D V K V T P G T K H V E
 GAGAGACCTAAAGTCCAAGAGATTCAAGCGGCGACCCCATGGCCACCGCTACTTGTGCGCCATCCATAGAGAGTAGTCCGAAGGCTG (360)
 E R P K V Q E I Q A A D P M A T A Y L C A I H R E L V R R L
 AAGGCCGCTCTGAACCGCTCTATACAGTGTGTTCGATATGAGCTCCGAGGATTTTGATGCTATCGTGGCCATGGGATGAAGTTGGGT (450)
 K A V L K P S I H V L F D M S S E D F D A I V G N G M K L G
 GACAAGGTGCTGGAACCGGACATCTCTCATTCGACAAGAGCCAGGACCAAGCCATGGCGTTACAGCGCTGATGCTGCTGAGGAGCTTG (540)
 D K V L E T D I S S F D S Q D Q A M A V T A L M L L R D L
 GGAGTAGAAGAAGCCTCTGACCTAATTGAGGCGCTTTTCGCGACATCACTTCTGCCACCTGCCACAGGCACAGATTTCAGTTT (630)
 G V E E D L L T L I E A S F G D I T S A H L P T G T R F Q F
 GGATCGATGATGAAGTCTGGACTTTTCTGACGCTGTCTGTAACACGCTGCTTAACATCACCATAGCTGCCGAGTTTACGGGAGCAG (720)
 G S M M K S G L F L T L F V N T L L N I T I A A R V L R E Q
 CTGGGTGATACAGGTGTGCGCGCTTTATCGGTGACGACACGTAATCACCAGGATAGTGTCTGACGACATGATGGTGGCCAGGTGCCA (810)
 L A D T R C A A F I G D D N V I T G V V S D D M M V A R C A
 TCCTGGCTGAACATGGAGGTGAAGATCATGACATGGAAATGGCAACATGAGTCTTATTTTGTGGCGGCTTCTGTTACTCGACAC (900)
 S W L N M E V K I M D M E I G N M S P Y F C G G F L L D T
 GTAACAGGCCTGTAAGCGAGTGTGCGACCTGTAAACCGCTGATGAAGATGGGAAACCGGCCCTGAACGATCCAGAAACGGAGCTG (990)
 V T G T V S R V S D P V K R L M K M G K F A L N D P E T D V
 GACAGATGCCCGCCTCTGCGCGAAGTGAAGAGTGGTACAGAGTGGGATTCAGTGGCCACTGCAGGTGGCTGCCGCCACACGCTAT (1080)
 D R C R A L R E E V E S N Y R V G I Q W P L Q V A A A T R Y
 GGCGTGAACCACTCTGCGCCCAATGGCGATGGCCACGCTCGCCAGGACTTGAGATCGTACCTGGCGCGCGAGGGAGTACGTA (1170)
 G V N H L P L A T M A M A T L A Q D L R S Y L G A R G E Y V
 TCCTCTACGTCTAACCITTAATTTTCTCATCTACTTCCAAACATCATGTTTCCCATGCAATTCACCAACTCAGCTTACGCCAG (1259)
 S L Y V *
 M F P M Q F T N S A Y R Q (13)
 ATGGAGCCCATGTTTGCACCGGTTCCCGAGGACAGTACAGCGTACCGGCGCGCACTAAGCGCGCGAGGAGCCGCAAGTCCGCAAC (1349)
 M E P M F A P G S R G Q V Q P Y R P R T K R R Q E P Q V G N (43)
 GCGCGATTACTGCGCTCGCGAACCAGATGAGTGGCTCCAGTTGACAGTACGCTGCACTTCCCGCCAGGCGAGGGTGGACCGCGTGGG (1439)
 A A I T A L A N Q M S A L Q L Q V A G L A G Q A R V D R R G (73)
 CCAAGAGCTGTTGAGAAGCAAGCAGAAGAAGAAGTCTTCCAAACGAGAGAAACCCAAAGAGAAGAAGCAAAACCAACAG (1529)
 P R V L Q K N E Q K K K N S S N G E K P K E K K K K Q K Q (103)
 GAGAAGAAGGGAAGCGGTGCGAAGTCAAGAAGTCAAGAACTAGGAACCGACCCGGAAGGAGGTAAAGATCTCCGTAAAGTGTGCCGACAG (1619)
 E K K G S G G E K V K K T R N R P G K E V R I S V K C A R Q (133)
 AGCACCTTCCCGTGTACCAAGAGGTGCTATATCCGCTACGCTGCTGCTGATTGATCTCGCGTATTCAGCGCGCACAGTGAAGGT (1709)
 S T F P V Y H E G A I S G Y A V L I G S R V F K P A H V K G (163)
 AAGATGACACCCCTGAACTGGCAGATCAAGTTCAGGTGCGCGAGGACATGGACCTCGAAGCAGCTGCGTACCCGAAGAGCATGCCA (1799)
 K I D H P E L A D I K F Q V A E D M D L E A A A Y P K S M R (193)
 GACCAAGCGGTGAACACGACCATGATGGACAGGTGTACAACTGGGAGTATGGCACTATCAGAGTGGAGGATAATGTCATAATCGAC (1889)
 D Q A A R P A T M M D R V Y N W E Y G T I R V E D N V I I D (223)
 GCAAGCGGTAGCGGCAAGCCGGTGCAGTGGCAGGGCTATCCGACAACTCGGGAAGGTTGTTGGTATTGTCCTCGSAGGAGGACCC (1979)
 A S G R G K P G D S G R A I T D N S G K V V G I V L G G G P (253)
 GATGGCAGCGCACACGCTCTCGGTGATAGTTTTCGACAGAAGATGAAGGCTAGGAGATCGCTACAGTGGATGACATACCTTGGACA (2069)
 D G R R T R L S V I G F D K K M K A R E I A Y S D A I P W T (1)
 CGCGCTCCGCGCTCTGCTGCTGCTATGTTTATGCTGACCTACAATTCCAAACCTTCGATTGCTCCAAACCGTCTGCGCAGGAC (2159)
 R A P A L L L L P M V I V C T Y N S N T F D C S K P S C Q D (31)
 TGCTGCATTACTGCTGAACACGAGAGGCCATGACCATGCTGAAGCAATCTGAACGACCGCACTACTGGACCTACTCATTGCTGTC (2249)
 C C I T A E P E K A M T M L K D N L N D P N Y W D L L I A V (61)
 ACCACCTGTGGCTCGCGCCGAGAAAGAGGGCTGTGCTACGTCGCTGCGCGCTTTTACGACACACAGATCTCGCGCCCAACGACGCT (2339)
 T T C G S A R R K R A V S T S P A A F Y D T Q I L A A R A A (20)
 GCCTCCCATACAGGCGTACTGCGCCGATTGTGACGGAACAGCGTGTATCTCGCGATAGCCATCGACGAGGTGGTGGACGAGTGGCAGC (2429)
 A S P Y R A Y C P D C D G T A C I S P I A I D E V V S S G S (50)
 GACCACGCTCTCGCATGCGGGTGGTTCTCAATCGGAGTGAACGCTAAGGGTGGTGGCGGCGGTGAGACCTCTCTGCGATACCTGGGA (2519)
 D H V L R M R V G S Q S G V T A K G G A A G E T S L R Y L G (80)

FIG. 2. The complete translated nucleotide sequence of the SPDV structural region and the C-terminus of nsP4. The nucleotide sequence is numbered from the 5' terminus. Amino acids are numbered from the start of each protein. This nucleotide sequence was submitted to the GenBank and has been assigned Accession No. AJ012631.

AGGACGGGAGGTTTCACGCCGACAAACACGCGACTCGTGGTGGCGCAGCTGCAAGTGCGACGTGCTGCAGGCCACTGSCCRCTAC (2609)
 R D G K V R A A D N T R L V V R T T A K C D V L Q A T G H Y (110)
 ATCTGGCCAACTGCCAGTGGGCGAGGCTAACGGTGGCGCCACACTGGATGGCACCAGGCTCAATGCCACCAAGTTTTCGAACAC (2699)
 I L A N C P V G Q S L T V A A T L D G T R H Q C T T V P F R H (140)
 CAAGTAACGGAGAGTTTACCCAGAGAACGACGAGGCCACCATCTGTCCGACATGACCAAGAAATGCACAGATTTTCCACTACACCA (2789)
 Q V T E K F T R E R S K G H H L S D M T K K C T R F S T T P (170)
 AAAAGTCCGCCCTCTACCTCGTGTATGTATGACGCTCTGCGGATTCTGTAGAGATTAGCACCGTGTACATGCAGCGACAGCCAG (2879)
 K K S A L Y L V D V Y D A L P I S V E I S T V V T C S D S Q (200)
 TGACAGTGGAGGTGCCACCTGGTACCACAGTGAATTCGACAAAGAAATGCAGAGCGCTGACTCGGCAACCGTCACTTTACACGCGAC (2969)
 C T V R V P P G T T V K F D K K C K S A D S A T V T F T S D (230)
 TCCAGAGCTTTAGGTGTGAGGAGCCAGTCTTACCGGCTGCCAGTATCACCCAGGCGAAGCCACCTCAGATCGGCAATGTTGCCATGC (3059)
 S Q T F T C E E P V L T A A S I T Q G K P H L R S A M L P S (260)
 GGAGGCAAGGAAGTGAAGCAAGGATCCCGTTCGGTTCGCCGCGAAGCCGCACTTGACAGTGTAGTGTAGCCCACTGCCCTCGATC (3149)
 G G K E V K A R I P F P P P P E T A T C R V S V A P L P S I (290)
 ACCTACGAGGAAGCGATGCTCTGCTAGCGGCTACCGCAAAATACCGTGTCTGCTGCTAACCACACGAACTTGGTTTCCATAGCAACGCC (3239)
 T Y E E S D V L L A G T A K Y P V L L T T R N L G F H S N A (320)
 ACATCGAATGGATCCAGGCGAAGTACCTGCGCCGATCCCGGTCAAGGATCGAGTAACTGGGAAACCAACGCGCCGATG (3329)
 T S E W I Q G K Y L R R I P V T P Q G I E L T W G N N A P H (350)
 CACTTTTGGTCACTCGTCAAGTACGCAACCGGAGCGCTGATGCTACCCCTGGGAATCTCTGTTGTACCAACCAAGCAACATCCAGAG (3419)
 H F N S S V R Y A S G D A D A Y P W E L L V Y H T K H H P E (380)
 TACCGTGGGCGTTTGTAGAGTTGCATGCGGCTGCTGGCTATCGCAGCGTGCATGTTTGGCGCGCATGCAGCGGTTGCGTACTCT (3509)
 Y A W A F V G V A C G L L A I A A C M F A C A C S R V R Y S (410)
 CTGGTCCGCAACGCTTCACTCGAAGCCACCAATGACCGCACTGACTGCGAGCACTGTGTTGCATACAGGGGCTCGCGCGGACCA (3599)
 L V A N T F N S N P P P L T A L T A A L C C I P G A R A D Q (2)
 CUCTACTGGACATCAITGCTACTTTTGTAGGGTAAAGGGTGGTCAAGCGCTGGTCTCTTGTGCTATGTACAGAGCTGCAAGAC (3689)
 P Y L D I I A Y F L G V R G W S A L L V I L A Y V Q S C K S (32)
 TACGAACACACCGTGGTGGTCCCAATGGATCCAGAGCCCGCTGCTACGAAGCAGTGATAAACCGGAATGGGTATGATCCATGGAAGCTG (3779)
 Y E H T V V V P M D P R A P S Y E A V I N R N G Y D P L K L (30)
 ACCATCTCAATGAATTTACCGTCACTCTACCAACTACCGCTCTGGAATATTGGACCTGCGCAGGAGTCCCATCTGCGAGCCGCCCAT (3869)
 T I S V N F T V I S P T T A L E Y W T C A G V P I V E P P H (60)
 GTGGCGCTGTCACGTCGGTGTCTCTGCGCTCTGACCTCTCTACGCTGCATGCGTTTACTGGCAAGCTGTCTCGAGCTGCACCTCGAT (3959)
 V G C C T S V S C P S D L S T L H A F T G K A V S D V H C D (90)
 GTGCACACAACTGTACCCCTGTGTTGTGGGGCGGGCTCACTGCTTCTGTTCACAGGAAATACAGAGTCAAGCGTGTGGCAGCCACC (4049)
 V H T N V Y P L L W G A A H C F C S T E N T Q V S A V A A T (120)
 GTTCTGAGTTCTGTGCGCAGGACTCAGAGCGTGGCAAGCGTTCAAGTACACAGAGCTCACTCAGCGTCAAGGCTGCTGCTGACGCTT (4139)
 V S E F C A Q D S E R A E A F S V H S S S V T A E V L V T L (150)
 GGTGAAGTGGTGAAGCGAGTCCAGTGTACGTTGAGTGGAGCGGTAACTCAGCCAGGCGCACTGACCTCAAGATCTGGCTGGACCAATAA (4229)
 G E V V T A V H V Y V D G V T S A R G T D L K I V A G P I T (180)
 ACCGACTACTCCCATTCGATCGCAAGTAGTCCGCTCGCGAGAGGTCTATACTATGACTGCGCTCTTACCGGGCTGGCGGACCA (4319)
 T D Y S P P D R K V V R I G E E V Y N Y D W P P Y G A G R P (210)
 GGCACATTCGGAGACATCAAGCTAGGTCAACCACTATGTCAAAACCAACGATCTGTATGGGACATCGGAATGAAGTACTGCGAGCG (4409)
 G T F G D I Q A R S T H Y V K P N D L Y G D I G I E V L Q P (240)
 ACTAAGCAGCAGTACATGTGCGCTTACAGGTATACGCTCTGCGGTTACTGCGTGGCGTGCAGGACGCTCCGAAACCACTAGTGTACA (4499)
 T N D H V H V A Y T Y T T S G L L R W L Q D A P K P L S V T (270)
 GCACCGCAGGTTGTAGATCAGTGCATATCCGCTCTCGGCTCTGATGTTGGGTTGGTGGCGTCCCATGTCCATCAACATTCGGGAC (4589)
 A P H G C K I S A N P L L A L D C G V G A V P M S I N I P D (300)
 GCGAAGTTTACCGCAAAATTAAAGGATCCGAAACCATCGGCTCGAAATGCGTGGTGGACAGCTGCGAGTADGGGGTGGACTACGGGGC (4679)
 A X F T R K L K D P K P S A L K C V V D S C E T G V D Y G G (330)
 GCGCGACGATCACTACGAGGCGCACGAGCGCGGAACTCGGGATTCTTCCGTGACACGAGGAGTCCCGCTGAGAATCATCGTGGTT (4769)
 A A T I T Y E G H E A G K C G I H S L T P G V P L R T S V V (360)
 GAAGTGGTGTGCGCGCAATACCGTCAAAACGACCTTCTCTACCCACGCGCGAGGTTGCACTCGAGTAGAGATCTGTGCGCAATA (4859)
 E V V A G A N T V K T T F S S P T P E V A L E V E I C S A I (390)
 GTGAAGTGGCTGGTGTGCTCCACCGAAGGAATGTGTCGCAACGAGGCTCGCCATGGCAGGACCTGGAGGCTACATCTCC (4949)
 V K C A G E C T P P K E H V V A T R P R H Q S D P G G Y I (420)
 GGGCGCGCAATGCGTGGCGGAGGAGTGTAGGAGCGCTAGTGTCTCTTCTTATCTTCTGCGTCACTCTAGTGGTGAAGAAG (5039)
 G P A M R W A G G I V G T L V V L F L I L A V I Y C V V K K (450)
 TCGCGCTCCAAAGAAATCCGATAGTCAAGAGCTAAATTCGGTATACAAATGTCTACTAGGAGCCCATCCGATCCACAGGGAGTAGG (5129)
 C R S K R I R I V K S * (461)
 ATGAGTCACTATGGTTTTAAATTTCAATAC - poly(A) (5163)

FIG. 2—Continued

TABLE 1
Size in Amino Acids of SPDV Structural Proteins Compared
to Nine Other Alphaviruses

	Capsid	E3	E2	6K	E1
SPDV	282	71	438	32	461
SIN	264	64	423	55	439
SFV	267	66	422	60	438
VEE	275	59	423	55	442
EEE	259	63	420	56	441
WEE	259	60	423	55	439
RRV	270	63	422	60	438
ONN	260	64	423	61	439
AURA	267	61	424	54	438
BFV	253	68	421	58	439

Note. Predictions were made on the basis of conserved cleavage sites between the structural proteins.

served hydrophobic stretch in its N-terminus believed to have a role in membrane fusion (7, 25).

Comparison of SPDV Noncoding Nucleotide Sequences with Those of Other Alphaviruses. The nucleotide sequence of the 26S RNA junction region of SPDV was compared with other alphaviruses (Fig. 3A). This region, conserved among alphaviruses, contains the C-terminus of nsP4 and the beginning of the 26S RNA and has a role as promoter for the transcription of the sub-genomic 26S RNA (20, 15). SPDV is similar in this region but not identical, showing nine differences to SIN and five differences to EEEV and WEEV over a 24-nucleotide stretch. SPDV also differs at three residues which are conserved in 10 other alphaviruses.

The 3'-nontranslated region of alphaviruses is believed to play an important role in virus replication. The 3'-nontranslated region of SPDV is one of the shortest reported for an alphavirus at 91 nucleotides, with Pixuna virus being the shortest at 77 nucleotides and Bebaru virus being the longest at 609 nucleotides (23). Whereas

most alphaviruses have two or three copies of repeat elements in their 3'-nontranslated regions (21), SPDV lacks repeat sequence elements and shows little homology to other alphaviruses in this region prior to the 19-nucleotide terminal sequence.

The last 19 nucleotides prior to the poly(A) tract are highly conserved among alphaviruses (21, 15). However, SPDV shows significant divergence in this region (Fig. 3B), in that it contains a nine-nucleotide stretch not seen in any of other sequenced alphaviruses. When this insertion and other small deletions are taken into consideration, SPDV shares nucleotide identity with other alphaviruses at 10-14 positions. Also, the 3'-terminal C residue prior to the poly(A) tract, which is believed to be necessary for replication of SIN virus (13), is present in the SPDV genome.

Discussion. In this paper we report the cloning, sequencing, and sequence analysis of the 3'-terminal 5.2-kb region of the genome of SPDV, a virus that infects and causes economically important disease in farmed Atlantic salmon. On the basis of the genomic organization of this region and protein homologies observed between the structural proteins of SPDV and other well-characterized alphaviruses, we have concluded that SPDV can be classified as an alphavirus, the first alphavirus reported in fish. On the basis that PD of farmed Atlantic salmon and sleeping disease (SD) of freshwater-reared rainbow trout, *Oncorhynchus mykiss*, share similar histopathology and that acquired cross-protection has been demonstrated, Boucher and Baudin Laurencin (7) have reported that the two diseases are caused by similar or identical agents. A comparative nucleotide sequence analysis will be useful for determining whether the virus associated with SD is an additional fish alphavirus and, if so, for establishing the extent of its similarity to SPDV.

The biological properties of SPDV are very different from those of other alphaviruses. For example, most

TABLE 2
Percentage Identities between the Structural Proteins Showing the Relationship of SPDV to Other Alphaviruses

	SPDV	SIN	SFV	WEE	EEE	VEE	ONN	RRV	AURA	BFV
SPDV	—	32	32	32	33	32	33	32	33	33
SIN		—	46	66	49	47	43	47	61	46
SFV			—	45	48	46	60	73	45	56
WEE				—	56	51	43	44	56	44
EEE					—	56	47	46	47	47
VEE						—	46	46	47	43
ONN							—	59	44	52
RRV								—	45	53
AURA									—	45
BFV										—

Note. Amino acid sequence information was obtained from Accession Nos. J02363(SIN), J02361(SFV), J04332(VEE), J03854(WEE), X63135(EEEV), M20162(RRV), M20303(ONN), S78478(Aura), U73744(BFV), and AJ012631 (SPDV).

A		B	
SPDV	C C C U C U A C G U C U A A C C U U A A U A U U	SPDV	U A U O G G U U U U A A A A U U U C A A U A C(A)n
VEE	- U - - - - - G - - - - - A - - - G G A	SFV	A-----U---- ----UUU C-(A)n
EEE	- - - - - G - G - - - A - - - G G	RRV	A--U-UUU--- ----UUUA-(A)n
WEE	- - - - - G - G - - - A - - - G G	VEE	AU-U- U----- --UUU -(A)n
SIN	A U - - - - - G U G G U - - - A - - - G -	SIN	AU-U- U----- --C-UUU -(A)n
MID	A - - - - - G - G G U - - - A - - - G -	WEE	AU-U- U----- --A-UUU -(A)n
RRV	A - - - - - G - G G U - - - A - - - G A	HJ	AU-U- U----- --A-UUU -(A)n
SFV	A - - - - - G - G G U - - - A - - - U G G	EEE	AU-U- U----- ----UUU -(A)n
ONN	- - U - G - - - G - G G - - - A - - - G G	ONN	AU-U- U----- ----UUU -(A)n
AURA	A - - - - - G U G G U - - - A - - - G A	MID	CCA-U- G----- ----UUC -(A)n
BFV	A U - - - - - G U G G U - - - A - - - G -	AURA	AU-U- U----- ----UUU -(A)n
		BFV	A--U- UU----- ----UUUA-(A)n

FIG. 3. (A) Comparison of the 26S junction region of SPDV and 10 other alphaviruses. Dashes indicate identical nucleotides. Those nucleotides found only in SPDV are shown in bold. (B) Comparison of the 19-nucleotide conserved region at the 3' terminus. Gaps have been introduced for alignment of this region of SPDV to other alphaviruses. Dashes indicate identical nucleotides.

previously described alphaviruses are arthropod or insect transmitted, generally by mosquitoes (2). Given that most of the fish farms experiencing SPDV-related disease are located in Northern Europe, for climatic reasons, it is extremely unlikely that mosquito-like insects play active roles in the transmission of this virus. However, the role that the ectoparasitic sea lice species (*Lepeophtheirus salmonis*) may play in the transmission of SPDV remains to be determined. Due in part to the fact that they can replicate in insect as well as animal hosts, some of the previously described alphaviruses, for example SFV, have been reported to infect a very broad range of cultured cells, including those of mammalian, avian, amphibian, reptilian, and insect origin (9, 14, 4, 16). In addition, research performed in this laboratory has shown that SFV can produce infectious virus following inoculation of CHSE-214 cells incubated at 25°C, although productive infections did not occur when this virus was grown at 20°C or below (Phenix and Todd, unpublished results). In contrast, the cell culture growth of SPDV virus is restricted to a very narrow range of salmonid fish cell lines maintained at temperatures close to 15°C.

Given the differences in the host species infected, it is not surprising that the levels of structural protein homology that exist between SPDV and other alphaviruses (32–33%) are less than those observed when previously characterized alphaviruses were compared (Table 2). Although, as might be expected for the nonstructural protein region, pairwise comparisons involving the C-terminal 394 amino acids of nsP4 show higher levels of identity (59–62%) between SPDV and other alphaviruses, these values are substantially less than those identities (75–83%) determined when other alphaviruses are compared over the same region. Our finding that the protein homology levels that SPDV shares with each of the other alphaviruses is relatively uniform (32–33%) supports the view that SPDV does not share a particularly close relationship with any of the previously characterized alphaviruses. Sequencing of the entire genome, which was not within the remit of this investigation, will be required

before the exact phylogenetic relationship of SPDV to other alphaviruses can be established.

Our sequence analysis has identified clear molecular differences between SPDV and previously sequenced alphaviruses. At 91 nucleotides the 3'-nontranslated region of SPDV is one of the shortest observed for any alphavirus. This region lacks repeat sequence elements that are detected in other alphaviruses and with which a possible role in virus RNA translation has been identified (28). In addition, SPDV shows the greatest level of nucleotide variation within the 19-nucleotide terminal sequence. This region, which is highly conserved among alphaviruses, is reported to be involved in the virus RNA replication (13). Extensive molecular investigations will be required before the significance of such differences can be determined.

Materials and Methods. Cells and virus. The F93125 isolate of SPDV used in this investigation was grown in Chinook salmon embryo (CHSE-214) cells as previously described (19). For virus purification purposes, monolayer cultures of CHSE-214 grown to ~80% confluence in 75-cm² flasks were infected with 1 ml virus to give a m.o.i. of ~1. After 1 h adsorption, an additional 14 ml supplemented Eagle's minimum essential medium (MEM) was introduced to each flask (19). The virus-infected flasks were incubated at 15°C for 7 or 8 days, when virus-induced cytopathic effect was evident, and the supernatant was collected.

Virus purification. The supernatant (typically 500 ml) from virus-infected cells was clarified at 1500 *g* for 20 min. Polyethylene glycol (PEG) and NaCl were added to give final concentrations of 6 and 2.2%, respectively. Following overnight incubation at 4°C the PEG precipitate was collected by centrifugation for 90 min at 1500 *g*. The resultant pellet was resuspended in PBS (1–2 ml) and the crude virus suspension was fractionated by equilibrium density centrifugation using 11-ml gradients (20–60% w/w in TNE) of sucrose. After centrifugation for 18 h at 50,000 *g* at 4°C, 1-ml fractions were collected from the bottom of the gradient. Fractions containing

virus were identified by immunoblotting using an SPDV-specific mouse monoclonal antibody (Welsh, unpublished results).

Production of SPDV cDNA Clones. Viral RNA was extracted from gradient-purified SPDV and virus-infected cells using RNA isolator (Genosys) and stored as ethanol precipitates. A cDNA library was first made by random priming with RNA extracted from gradient-purified virus. This library consisted of clones containing inserts (250–500 bp) in the vector pUC18 (Sureclone ligation kit, Pharmacia). Clones were selected randomly from the library and following sequencing and analysis using the BLAST program (University of Wisconsin, Genetics Computer Group) were mapped to the 'alphavirus' genome. The sequences of three such clones, N11, N38, and N50, which mapped to approximately 5.2, 2.7, and 1.1 kb from the putative 3' terminus of the 'alphavirus' genome, were exploited to design oligonucleotide primers, which were used, in reverse transcription-polymerase chain reactions (RT-PCR), to amplify three overlapping fragments encompassing the 5.2-kb region at the 3' terminus of the SPDV genome. The incorporation of *NotI* sites into the primers facilitated the restriction ligation of two of these fragments into the *NotI* site of the vector pBluescript (Stratagene) to produce clones pBS1150 and pBS5038. For pBS1150 first strand synthesis was carried out using the primer N50revNot (5'-GCATGCGGCCGCGATGCGGAG-GACGTGGTCGCT-3') followed by PCR using this primer and the forward primer N11forNot (5'-GCATGCGGCCGCGCCGACTATGGACTCAGCGGCA-3'). For pBS5038 first strand synthesis was carried out using the primer N38revNot (5'-GCATGCGGCCGCGCTGAACGCTTCGGC-ACGCTC-3') followed by PCR using this primer and the forward primer N50forNot (5'-GCATGCGGCCGCGAGC-GACCAAGTCTCCGCAT-3'). Both PCRs were carried out using the Expand Long Template PCR System (Boehringer Mannheim) at 94°C for 30 s, 60°C for 30 s, 68°C for 2 min. Clone pBS38R was produced using 3' RACE (6). The reaction was performed using a 5'/3' RACE kit (Boehringer Mannheim) with some modifications. Thus, RNA from gradient-purified virus was independently subjected to first strand cDNA synthesis using an oligo d(T) anchor primer (5'-GACCACGCGTATCGAT-GTCGAC(T)₁₆-3'). Resultant cDNAs were amplified by PCR at 94°C for 30 s, 60°C for 30 s, 68°C for 1 min using the forward primer designed from N38, containing a *NotI* restriction site (N38forNot (5'-GCATGCGGCCGCTCA-GAGCGTGCCGAAGCGTTC-3')) and the modified anchor primer (5'-GCATGCGGCCGCGGACCAAGCGTATCGATGT-CGA-3') containing a *NotI* site allowing cloning into pBluescript.

Sequencing of SPDV cDNA Clones. Cycle sequencing was performed using the ABI PRISM dye terminator ready reaction kit on purified plasmid DNA following the manufacturer's protocol (Perkin-Elmer Cetus). Electro-

pherograms were interpreted using the Sequence Navigator software (Perkin-Elmer Cetus). The sequencing strategy adopted for SPDV is shown in Fig. 1. Sequencing of the clones pBS1150, pBS5038, and pBS38R was initially carried out using primers specific to the T7 and T3 sequences within the plasmid vector and then by a primer walking strategy. The three constructs were sequenced on both strands more than twice. Sequence comparisons were carried out using the GCG package, Wisconsin package version 8.1-UNIX, Genetics Computer Group Inc.

ACKNOWLEDGMENTS

Financial support for this research was provided by Intervet International, Boxmeer, The Netherlands. The authors acknowledge the contribution made to this work by Drs. Brian Meehan (The Queen's University of Belfast) and John Claessens (Intervet International, Boxmeer, The Netherlands).

REFERENCES

1. Boucher, P., and Baudin Laurencin, F. (1996). Sleeping disease and pancreas disease: Comparative histopathology and acquired cross protection. *J. Fish Dis.* 19, 303–310.
2. Chamberlain, R. W. (1980). Epidemiology of arthropod-borne togaviruses: The role of arthropods as hosts and vectors and of vertebrate hosts in natural transmission cycles. In 'The Togaviruses' (R. W. Schlesinger, Ed.), pp. 175–227. Academic Press, New York.
3. Choi, H.-H., Tong, L., Minor, W., Dumas, P., Boege, U., Grossman, M. G., and Wengler, G. (1991). Structure of Sindbis virus core protein reveals a chymotrypsin-like serine protease and the organisation of the virion. *Nature* 354, 37–43.
4. Clark, H. F., Cohen, M. M., and Lunger, P. D. (1973). Comparative characterisation of a C-type virus-producing cell line (VSW) and a virus-free cell line (VH2) from *Viperia russelli*. *J. Natl. Cancer Inst.* 51, 645–657.
5. Forsell, K., Suomalainen, M., and Garoff, H. (1995). Structure-function relation of the NH₂-terminal domain of the Semliki Forest virus capsid protein. *J. Virol.* 69, 1556–1563.
6. Frohmann, M. A., Dush, M. K., and Martin, G. R. (1988). Rapid production of full-length cDNAs from rare transcripts using a single gene-specific oligonucleotide primer. *Proc. Natl. Acad. Sci. USA* 85, 8998–9002.
7. Garoff, H., Frischaut, A.-M., Simons, K., Lahrach, H., and Delius, H. (1989b). Nucleotide sequence of cDNA coding for Semliki Forest virus membrane glycoproteins. *Nature* 288, 236–241.
8. Geigenmüller-Guirke, U., Nitschko, H., and Schlesinger, S. (1993). Deletion analysis of the capsid protein of Sindbis virus: Identification of the RNA binding region. *J. Virol.* 67, 1620–1626.
9. Griffin, D. E. (1986). Alphavirus pathogenesis and immunity. In 'The Togaviridae and Flaviviridae' (S. Schlesinger and M. J. Schlesinger, Eds.), pp. 209–250. Plenum, New York.
10. Hahn, C. S., and Strauss, J. H. (1990). Site directed mutagenesis of the proposed catalytic amino acids of the Sindbis virus capsid protein autoprotease. *J. Virol.* 64, 3069–3073.
11. Kent, M. L., and Elston, R. A. (1987). Pancreas disease in pen-reared Atlantic salmon in North America. *Bull. Eur. Assoc. Fish. Pathol.* 7, 29–31.
12. Kerr, P. J., Fitzgerald, S., Tregear, G. W., Dalgarno, L., and Weir, R. C. (1992). Characterisation of a major neutralization domain of Ross River virus using anti-viral and anti-peptide antibodies. *Virology* 187, 338–342.

13. Kuhn, R. J., Hong, Z., and Strauss, J. H. (1990). Mutagenesis of the 3' nontranslated region of Sindbis virus RNA. *J. Virol.* 64, 1465-1476.
14. Leake, C. J., Varma, M. G. R., and Pudney, M. (1977). Cytopathetic effect and plaque formation by Arboviruses in a continuous cell line (XTC-2) from the toad *Xenopus laevis*. *J. Gen. Virol.* 35, 335-339.
15. Levis, R., Weiss, B. G., Tsiang, M., Huang, H., and Schlesinger, S. (1986). Deletion mapping of Sindbis virus DI RNAs derived from cDNAs defines the sequences essential for replication and packaging. *Cell* 44, 137-145.
16. L'Héritier, P. (1977). About the infection of *Drosophila* female germ line cells by Sigma viruses. *Ann. Microbiol.* 128A, 119-131.
17. McLoughlin, M. F., Nelson, R. T., Rowley, H. M., Cox, D. I., and Grant, A. N. (1996). Experimental pancreas disease in Atlantic salmon *Salmo salar* post-smolts induced by salmon pancreatic disease virus (SPDV). *Dis. Aquatic Organisms* 26, 117-124.
18. Munro, A. L. S., Ellis, A. E., McVicar, A. H., McLay, H. A., and Needham, E. A. (1984). An exocrine pancreas disease of farmed Atlantic salmon in Scotland. *Helgo. Meeresunters* 37, 571-586.
19. Nelson, R. T., McLoughlin, M. F., Rowley, H. M., Platten, M. A., and McCormick, J. I. (1995). Isolation of a toga-like virus from farmed Atlantic salmon *Salmo salar* with pancreas disease. *Dis. Aquatic Organisms* 22, 25-32.
20. Ou, J.-H., Rice, C. M., Dalgarno, L., Strauss, E. G., and Strauss, J. H. (1982a). Sequence studies of several alphavirus genomic RNAs in the region containing the start of the subgenomic RNA. *Proc. Natl. Acad. Sci. USA* 79, 5235-5239.
21. Ou, J.-H., Trent, D. W., and Strauss, J. H. (1982b). The 3' non-coding regions of alphavirus RNAs contain repeating sequences. *J. Mol. Biol.* 156, 719-730.
22. Owen, K. E., and Kuhn, R. J. (1996). Identification of a region in the Sindbis virus nucleocapsid protein that is involved in specificity of RNA encapsidation. *J. Virol.* 70, 2757-2763.
23. Pfeffer, M., Kinney, R. M., and Kaaden, O.-R. (1998). The alphavirus 3'-nontranslated region: Size heterogeneity and arrangement of repeated sequence elements. *Virology* 240, 100-108.
24. Poppe, T., Rimstad, E., and Hyllseth, B. (1989). Pancreas disease (*Salmo salar*) post-smolts infected with infectious pancreatic necrosis virus (IPNV). *Bull. Eur. Assoc. Fish. Pathol.* 9, 83-85.
25. Rice, C. M., and Strauss, J. H. (1981). Nucleotide sequence of the 26S mRNA of Sindbis virus and deduced sequence of the encoded virus structural proteins. *Proc. Natl. Acad. Sci. USA* 78, 2062-2066.
26. Rümenapf, T., Strauss, E. G., and Strauss, J. H. (1995). Aura virus is a new world representative of Sindbis-like viruses. *Virology* 208, 621-633.
27. Strauss, E. G., Stec, D. S., Schmaljohn, A. L., and Strauss, J. H. (1991). Identification of antigenically important domains in the glycoproteins of Sindbis virus by analysis of antibody escape variants. *J. Virol.* 65, 4654-4664.
28. Strauss, J. H., and Strauss, E. G. (1994). The alphaviruses: Gene expression, replication, and evolution. *Microbiol. Rev.* 58, 491-562.
29. Wheatley, S. B. (1994). "Epidemiological Investigations of Atlantic Salmon Production Using a Computerised Database System." Ph.D. thesis, The Queen's University of Belfast.